TQ RESEARCH & POLICY

IMPROVING INSTRUCTION THROUGH EFFECTIVE TEACHER EVALUATION: OPTIONS FOR STATES AND DISTRICTS

February 2008

Carrie Mathers
Michelle Oliva
With Sabrina W. M. Laine, Ph.D.

The purpose of this Research and Policy Brief is to provide state and local policymakers with a comprehensive understanding of the measures used in teacher evaluation—their strengths, limitations, and current use in policy and practice.

This brief will underscore aspects of evaluation policies currently aligned with best practices as well as illuminate areas where policymakers may improve evaluation rules, regulations, and their implementation, thereby improving teacher instruction and student performance.

Contents

Page

TEACHER EVALUATION:
A LEVER FOR INSTRUCTIONAL IMPROVEMENT

The research clearly shows a critical link between effective teaching and students' academic achievement. In fact, a National Comprehensive Center for Teacher Quality 2007 synthesis of research concludes that although many studies point to outcomes that show some teachers contribute more to their students' academic growth than other teachers, almost no research can systematically explain the considerable variation in teachers' skills for promoting student learning (Goe, 2007). Pinpointing the skills that lead certain teachers to have a greater impact on student performance than others is a matter of great urgency in a country that struggles with educating all of its children equally. The growing interest in better understanding what constitutes effective teaching practice, coupled with its power to leverage educational improvement, presents a challenge and opportunity for policymakers to address how to efficiently and reliably measure teacher performance. The role of teacher evaluations has surfaced only recently as an underutilized resource that might hold promise as a tool to promote teacher professional growth and measure teacher effectiveness in the classroom.

When used appropriately, teacher evaluations should identify and measure the instructional strategies, professional behaviors, and delivery of content knowledge that affect student learning (Danielson & McGreal, 2000; Shinkfield & Stufflebeam, 1995). There are two types of evaluations—formative and summative. Formative evaluations are meant to provide teachers with feedback on how to improve performance and what types of professional development opportunities will enhance their practice. Summative evaluations are used to make a final decision on factors such as salary, tenure, personnel assignments, transfers, or dismissals (e.g., Barrett, 1986). Although both types of evaluations seek to measure performance, the formative evaluation identifies ways to improve performance and the summative evaluation determines whether the performance has improved sufficiently such that a teacher can remain in his or her current position and be rewarded for performance. While each type is valuable, neither type of evaluation can serve a teacher and school well on its own. Without formative feedback, a teacher may not be informed of "areas of weaknesses" so when the summative evaluation takes place, these "areas of weaknesses" may still exist. Similarly, ongoing formative evaluations without any consequences provide minimal incentives for teachers to act on the feedback.

When coupled, formative and summative evaluations can be powerful tools for informing decisions about teachers' professional development opportunities (e.g., Nolan & Hoover, 2005) as well as tenure (Brandt, Mathers, Oliva, Brown-Sims, & Hess, 2007). This combination is important because of the expense related to professional development delivery and the effect that professional development can have on teacher satisfaction and retention. Although districts are spending millions of dollars on professional development, oftentimes teachers report dissatisfaction with their experiences and attribute this dissatisfaction as a major factor when considering leaving a school (Parkes & Stevens, 2000). Using evaluation results to create and implement professional development plans may improve how current resources are being spent, send a message to teachers that their professional growth is valued, and decrease turnover rates.

The value of quality evaluation systems does not stop there. Administrators' use of evaluation results to make well-substantiated personnel decisions can have a direct effect on student learning outcomes. For example, Gordon, Kane, and Staiger (2006) posited that if the Los Angeles School District (whose data they analyzed) were to drop the bottom quartile of teachers in terms of their value-added

impact on student test scores in the first year of teaching, the district could raise overall student achievement by 14 percentile points over 12 years. In sum, using evaluation results to inform professional development and personnel decisions would yield a much greater return on taxpayers' investments in public education.

Given the potential value of using teacher evaluation to improve teacher satisfaction and student learning opportunities, several questions merit consideration: What do current teacher-evaluation systems look like? Are current evaluation systems aligned with what the research and expert guidance suggest? If the answer to the second question is no, how should they be improved? This Research and Policy Brief answers these questions by reviewing various teacher evaluation tools and assessing the strengths and weaknesses of each. It also provides policy options that can guide state and local processes and the application of evaluation results designed to support teacher instruction. To inform this discussion, this brief considers major findings from a recent teacher evaluation study conducted by REL Midwest (Brandt et al., 2007).

WHAT WE KNOW ABOUT TEACHER EVALUATION SYSTEMS

Current approaches to teacher evaluation vary in their scope and intent. To date, only three descriptive studies have examined teacher evaluation policies on a large scale (Brandt et al., 2007; Ellett & Garland, 1987; Loup, Garland, Ellett, & Rugutt, 1996). All three studies used a version of the Teacher Evaluation Practices Survey (TEPS) (Ellett & Garland, 1987) to collect information about districts' teacher evaluation policies and procedures. In addition, the National Council on Teacher Quality (2006) compiled a database of teacher contracts in the nation's 50 largest districts, which includes some information on teacher evaluation.

Ellett and Garland (1987) surveyed superintendents and collected teacher evaluation policies from the 100 largest school districts in the United States. Analysis of the districts' policy documents suggested that (1) teacher evaluations emphasized summative (e.g., dismissal, remediation) rather than formative (e.g., professional development) purposes; (2) most policies did not include requirements for establishing performance standards and evaluator training; (3) few districts permitted external or peer evaluations; and (4) superintendents tended to present their district policies more favorably than the independent reviewers of those policies.

A decade later, Loup et al. (1996) conducted a follow-up study to Ellett and Garland's work; however, rather than collecting the 100 largest school districts' policies, the researchers adapted the TEPS to measure superintendents' opinions about the effectiveness of their evaluation systems. Their TEPS results mirrored those from Ellett and Garland. Although a decade had passed, little had changed in regard to large districts' teacher evaluation policies. However, according to their reported opinions,

superintendents were not satisfied with the status quo. Many reported a need to revisit and revise their districts' existing evaluation tools and procedures (Loup et al., 1996).

While not focused exclusively on teacher evaluation policies, the database compiled by the National Council on Teacher Quality (2006) contains teacher contracts in the nation's 50 largest districts. An examination of policies contained in this database reveals a surprising lack of detail on local approaches to teacher evaluation.

Finally, the study released by REL Midwest in December 2007 collected teacher evaluation policies from a representative sample of districts in seven Midwestern states—Illinois, Indiana, Iowa, Michigan, Minnesota, Ohio, and Wisconsin (Brandt et. al, 2007). This study systematically describes local evaluation policies across a demographically diverse sample of districts. Its major findings are summarized in the accompanying sidebar on page 4. The complete study and information on the methodology can be accessed on the Regional Educational Laboratory Program website (http://ies.ed.gov/ncee/edlabs/projects/index.asp).

A Summary of the Major Findings From
Examining District Guidance to Schools on
Teacher Evaluation Policies in the Midwest Region

An analysis of the evaluation policies collected for the REL Midwest study on teacher evaluation policies (Brandt et al., 2007) indicated the following:

- Administrators (e.g., principals, vice principals) were most commonly charged with conducting evaluations.

- Only one half of the policies provided guidelines regarding when to conduct evaluations (e.g., fall and spring). Approximately two thirds of the policies detailed how often to evaluate teachers. Many of these policies required schools to differentiate evaluation frequency by teacher experience (i.e., probationary, tenured); however, policies rarely specified how often to evaluate teachers with previously unsatisfactory evaluations.

- A little more than one half of the policies identified the type of evaluation instrument to be used. The majority used summative rating scales to assess teacher performance. In almost all cases, the same evaluation was used—independent of the teacher's years of experience and subject area.

- Only one third of the policies detailed how to communicate the evaluation process and procedures to teachers. The most common methods of communication included teacher handbooks, group or one-on-one orientation, and contracts.

- One half of the policies required specific evaluation methods. The most common method was classroom observations (both scheduled and unannounced).

- Fewer than one third of the policies stated how to share the evaluation results with teachers. Most of the policies required teachers to sign off on the summative form after reviewing the evaluation.

- Almost one half of the policies included language about how the evaluation results should be used by administrators. The top four ways in which districts required evaluation results to be used (from most common to least common) were as follows: (1) to inform personnel decisions; (2) to make suggestions for teacher improvements; (3) to inform teacher professional development goals; and (4) to determine remediation or follow-up procedures (e.g., intensive improvement plan, coaching) for teachers with unsatisfactory evaluations.

- Just over one third of the policies identified teacher behaviors and characteristics to be evaluated. Most required the evaluation to measure content and pedagogical knowledge, classroom management skills (i.e., ability to engage students as well as maintain a positive learning environment), ability to effectively prepare a lesson, and the extent to which teachers fulfill their professional responsibilities. Only one half of the policies required an assessment of how well teachers use student progress to inform their teaching.

- Just more than one fourth of the policies identified the research and/or guidance informing their policy. The most commonly cited teacher evaluation model was the framework created by Charlotte Danielson (1996). A few districts referenced state standards.

- Fewer than one out of 10 policies required evaluator training.

CURRENT TEACHER EVALUATION TOOLS: STRENGTHS AND LIMITATIONS

Lesson plans, classroom observations (including video observations), self-assessments, portfolio assessments, student achievement data, and student work-sample reviews are often identified as common evaluation tools. The following paragraphs compare and contrast these diverse tools and the frequency with which district policies require them, with recommendations from research and expert opinion.

LESSON PLANS

Expert guidance often suggests the review of teachers' lesson plans as one evaluation method. Lesson plans are a window into a teacher's preparation to deliver content, scaffold the development of student skills, and manage the classroom learning environment. While some districts use rubrics to evaluate lesson plans (e.g., Denner, Salzman & Bangert, 2001), the REL Midwest study found that less than 4 percent of the 140 districts that submitted policies required lesson plans to be used as part of a teacher's evaluation (Brandt et al., 2007).

Strengths: One aspect of teaching correlated with student learning is the level of planning used to drive instruction (e.g., Stronge, 2007). Lesson plans are more likely to be positively related to improved student outcomes when plans are able to (1) link student learning objectives with teaching activities, (2) describe teaching practices to maintain students' attention, (3) align student learning objectives with the district and state standards, and (4) accommodate students with special needs (Stronge, 2007).

Limitations: It is important to remember that a lesson plan is indeed a "plan," and once it is implemented, the plan may need to be adjusted. The quality and appropriateness of the adjustments a teacher makes in the implementation of the plan in the classroom cannot be evaluated solely from the lesson-plan scoring rubric.

CLASSROOM OBSERVATIONS

Although teachers may be able to craft high-quality lesson plans, it is equally as important to link these plans with what occurs in the classroom. The classroom observation is the most commonly used tool for evaluating teachers. Of the 140 districts that submitted policies for the REL Midwest study on teacher evaluation policies, 41 (29 percent) suggested or required the use of formal observations, including scheduled observations (Brandt et al., 2007). The stark difference in the use of lesson plans (less than 4 percent) and classroom observations (29 percent) in the Midwest region suggests that evaluators rarely link planning to practice. Without the lesson plans, evaluators may be missing key information. For example, if student accommodations are needed for the lesson, it would be difficult for the evaluator to know if these accommodations are implemented appropriately without the lesson plan.

Strengths: Classroom observations capture information about teachers' instructional practices (Mujis, 2006). Observations can be used in formative and summative evaluations. When used in formative evaluations, the observation can track a teacher's growth and suggest needed professional development—the results of which can then be assessed in subsequent observations.

Limitations: Despite the frequent use of classroom observations for the purpose of evaluating teacher performance, this measure is not without its limitations. Poorly trained observers and inconsistent, brief observations can create biased results (Shannon, 1991; Shavelson, Webb, & Burstein, 1986). Research suggests that when observations occur more frequently, their reliability improves (Denner, Miller, Newsome, & Birdsong, 2002), and similarly, when observations are longer, their validity improves (Cronin & Capie, 1986).

Self-assessments

Reflection is a process in which teachers analyze their own instruction retrospectively. It can occur in a variety of ways: professional conversations with other teachers during grade or subject-area meetings (Uhlenbeck, Verloop, & Beijaard, 2002), preobservation and postobservation debriefings, the development of a portfolio, or an individual professional development plan. According to Brandt et al. (2007), only six of the participating districts required evaluations to determine how teachers use self-reflection to respond to student needs.

Strengths: Requiring reflection as part of an evaluation process may encourage teachers to continue to learn and grow throughout their careers (Uhlenbeck et al., 2002). To encourage reflection, some evaluation systems include videotaping teachers in the classroom. The videotaped class sessions may be rated as classroom observations, but these videotapes also allow teachers to review their performance so they can reflect and engage in in-depth conversations with their evaluators about the behaviors and practices observed.

Limitations: Reflection requires both time and a cultural norm that supports this type of evaluation practice in a school or district. When reflection is not typically used for evaluative purposes, making the time for teachers to engage

in this practice is a low priority for administrators (Peterson & Comeaux, 1990; Schon, 1983).

Portfolio Assessments

Portfolio assessments tend to comprise several pieces of evidence of teacher classroom performance, including lesson or unit plans, a video of classroom teaching, reflection and self-analysis of teaching practices, examples of student work, and examples of teacher feedback given to students (Andrejko, 1998). Portfolios are required in some states and districts, but they are less common than classroom observations. In the REL Midwest study, 13 out of 140 districts (9 percent) required portfolio assessments as part of their teacher evaluation system (Brandt et al., 2007).

Strengths: Teachers and administrators often favor the use of portfolios because they enable teachers to reflect on their own practice, allow evaluators to identify teachers' instructional strengths and weaknesses, and encourage ongoing professional growth (Attinello, Lare & Source, 2006; Tucker, Stronge, & Gareis, 2002). According to Danielson (1996), portfolios are useful evaluation tools because they allow evaluators to review nonclassroom aspects of instruction as well as provide teachers with opportunities to reflect on their teaching by reviewing documents contained in the portfolio. Portfolios also promote the active participation of teachers in the evaluation process (Attinello et al., 2006).

Limitations: Currently, there are no conclusive findings on the reliability of portfolio assessments as part of an objective teacher-evaluation system (Attinello et al., 2006). Existing research has raised questions about whether portfolios accurately reflect what occurs in classrooms and whether the process of developing a portfolio and being evaluated through that process leads to improvements in teaching practices (e.g., Attinello et al., 2006). The necessary time to develop and review a portfolio is another frequently cited concern (e.g., Attinello et al., 2006; Tucker et al., 2002).

STUDENT ACHIEVEMENT DATA

In addition to, or in place of, direct evaluations of teachers' characteristics and behaviors, some evaluation systems use standardized student test scores to assess the teacher's contributions to student learning. To isolate the effects of a teacher on student learning, such systems use statistical techniques and models to analyze changes in standardized test scores from one year to the next. Some examples of statistical models include the use of proficiency standards for measuring adequate yearly progress (AYP) of various student subgroups, the increasing use of value-added models, and the application of growth models that measure changes in student performance over time (longitudinally).

Although districts throughout the United States use these techniques, none of the 140 district policies collected as part of the REL Midwest study required student achievement data to be used as part of a teacher's evaluation (Brandt et al., 2007).

Strengths: The use of standardized student test scores enables schools to measure the impact that instruction is having on student performance and builds on an existing investment in student testing. While the quality of state and local assessments differ widely, the items on a well-developed standardized student assessment have been tested for issues of fairness and appropriateness through the application of various statistical models. Therefore, schools have an opportunity to examine the relationship between changes in student achievement gains, teachers, and schools (Braun, 2005). Recent case studies demonstrate how schools are taking advantage of this approach to enhance their teacher evaluations (e.g., Gallagher, 2004; Milanowski, 2004).

Limitations: Standardized student test scores measure only a portion of the curriculum and teachers' effects on learning (Berry, 2007). Most statistical models are not able to differentiate which elements of teaching relate to positive student achievement test outcomes. For example, Teacher A consistently improves students' fifth-grade reading scores; in sixth grade, however, the same group of students' reading scores are stagnant or decline in Teacher B's class. What is Teacher A doing that consistently and positively improves students' reading trajectories? Or is it something about Teacher B's behavior or something in the context of this particular classroom that is constraining Teacher B's practice? Moreover, as this example illustrates, teachers' value-added effects on test scores are meaningful only in relation to one another, rather than to established teaching proficiency criteria.

Confounding comparisons is an issue with statistical models, such as those used for AYP. It could be that one year's cohort consists of less prepared students and the following year's cohort (same grade, different students) consists of more motivated and better prepared students. Either way, they are not the same students, and the high performers will have less difficulty meeting proficiency standards than low-performing students.

A distinctly different concern with value-added models is that they depend on elaborate databases and data software that can link student and teacher data. Moreover, even with an adequate data infrastructure, not all teachers can be assessed using student test scores. Those who teach social studies, physical education, music, art, special education—as well as K–2 teachers and many middle and high school teachers—cannot be assessed using student test scores because not all are assigned a defined set of students in a classroom and not all students are tested every year or in every subject (e.g., social studies teachers).

STUDENT WORK-SAMPLE REVIEWS

An emerging view is that there may be alternative ways to measure the effect of instruction on student learning, including the analysis of student work samples (Mujis, 2006). This method is intended to provide a more insightful review of student learning results over time. Although district policies did not specify student work samples as part of the evaluation in the REL Midwest study, 22 districts' policies

required that the teacher evaluations contain components to gauge whether teachers examine their students' performance through measures such as assessment data (Brandt et al., 2007).

Strengths: Using student work samples as the basis for a review of teacher practice, one study found a large discrepancy between students' standardized reading scores and students' reading levels (Price & Schwabacher, 1993). This result suggests that student work samples may help to better identify which elements of teaching relate more directly to increased student learning than standardized test scores.

Limitations: One drawback to using student work samples in evaluations is that reviewing these samples can be time-consuming. In addition, the review of student work samples as a means of

evaluating teacher effectiveness is more prone to issues of validity and reliability than are achievement test items that have been validated for similar comparisons across different students in different schools answering similar test items. (Reliability and validity are discussed in the sidebar below.) To reduce subjectivity and address issues of reliability, experts should develop a research-informed scoring rubric that outlines criteria for rating student work samples. Those using the rubric should be trained so that the process is consistent across all student sample evaluations.

THE IMPORTANCE OF RELIABILITY AND VALIDITY IN TEACHER EVALUATION

An evaluation instrument is considered *reliable* if two or more evaluators use the same evaluation instrument and come to the same conclusion. For example, if a principal and a teacher leader evaluate Teacher A under similar conditions (e.g., same classroom, same students, and similar content being taught) and use the same evaluation instrument, then both should arrive at the same conclusions. One way to increase reliability is to ensure that the evaluation instrument has clearly defined, nonsubjective criteria that require minimal interpretation. This goal is accomplished by carefully developing evaluation instruments (e.g., pilot-testing the instruments before using them) and training observers (Mujis, 2006). Without these steps, the system collects data that cannot be transformed into meaningful information.

In addition to ensuring that evaluation measures are reliable, designers of teacher evaluation systems must ensure that evaluation tools are *valid*—that is to say, that the rubric or observation form assesses the teaching performance it was designed to measure. A first step in determining validity is for school staff to examine the proposed evaluation form to see whether "on its face" it seems like a good translation of teacher performance. Once there is staff consensus that the tool appears to accurately assess what it is designed to assess, that relationship must be tested. Developers must conduct several pilot trials with teachers and administrators to sharpen the instrument's language and process of implementation to ensure that what is being measured is clear and there is shared understanding of the district's definition of "excellent teaching performance." If the evaluation depends on student data, then in addition to criteria for teacher characteristics and behaviors, the definition should outline the desired improvements and changes in student behaviors, performance, and learning that "excellent teaching performance" is expected to produce. With adequate data, developers can descriptively and statistically demonstrate the link between teacher performance and student outcomes such that the excellent teaching performance being measured in fact produces the desired improvements in student behaviors, performance, and learning.

EVALUATION PROCESSES: REALITY VERSUS BEST PRACTICE

As mentioned in the discussion of evaluation tools, the validity and reliability of instruments designed to measure teacher performance are affected by the processes and procedures used to carry out teacher evaluations. This section compares and contrasts processes and procedures commonly used by districts with those recommended by research and expert opinion. For examples of evaluation innovations, see the sidebar on page 11.

WHO EVALUATES

Reality: Administrators (e.g., principals, vice principals) are the most common evaluators. According to the REL Midwest study, of the 140 Midwestern districts that provided policy and procedural documentation, 57 (41 percent) identified the position(s) responsible for conducting teacher evaluation; 44 of the 57 districts (77 percent) identified building administrators as the teacher evaluators (Brandt et al., 2007).

Recommended: Teachers highly regard evaluators with deep knowledge of curriculum, content, and instruction who can provide suggestions for improvement (e.g., Stiggans & Duke, 1988; Wise, Darling-Hammond, McLaughlin, & Bernstein, 1984). Therefore, multiple evaluators—peers who have an instructional

background, content knowledge, and experience teaching similar students—are a growing alternative to an administrator as the sole evaluator (e.g., Goldstein & Noguera, 2006).

FREQUENCY OF EVALUATION

Reality: Nontenured teachers often are evaluated twice a year, and tenured teachers once every three to five years unless they receive an unsatisfactory evaluation (Brandt et al., 2007; Sweeney & Manatt, 1986). An evaluation that captures one single point in time as interpreted by one evaluator, especially when compounded by the use of a weak rubric, ultimately is not the most valid way to measure teacher performance. Together, these shortcomings reduce the evaluator's ability to authentically measure the teacher's instruction and capture changes over time. As a result, these one-time, fuzzy snapshots fall short of gauging teachers' strengths and limitations. When this situation is the case, the school misses the opportunity to increase teacher growth and ultimately student achievement.

Recommended: Infrequent evaluations, particularly of tenured teachers, create missed opportunities to inform teaching practices and improve student learning. Both nontenured and tenured teachers should receive frequent evaluations. Although there is limited research on how often teachers should be evaluated, research using video observations of teachers as part of the evaluation suggests that four or

five observations as part of a single evaluation would be ideal (Blunk, 2007). However, additional research and guidance are needed to determine and confirm the optimal frequency of evaluations for both nontenured and tenured teachers.

TRAINING

Reality: Districts rarely require evaluators to be trained (Brandt et al., 2007; Loup et al., 1996). In the REL Midwest study, only 11 districts (8 percent) had written documentation detailing any form of training requirements for their evaluators (Brandt et al., 2007).

Recommended: Lack of training can threaten the reliability of the evaluation and the objectivity of the results. Not only do evaluators need a good understanding of what quality teaching is, but they also need to understand the evaluation rubric and the characteristics and behaviors it intends to measure. Without adequate training, observers may be unaware of the potential bias that they are introducing during their observations. If an observer has a preconceived expectation of a teacher or is overly influenced one way or another by the local school culture and context, the observation may be aligned with this expectation rather than the actual behaviors displayed by the teacher during the observation (Mujis, 2006).

COMMUNICATION

Reality: District policies do not always require teachers to be informed of the evaluation process or the potential implications. In the REL Midwest study, 45 of the districts (32 percent) had formal documentation requiring the evaluation policy to be communicated to teachers (Brandt et al., 2007).

Recommended: Systematic communication about the evaluation should occur with teachers prior to, during, and after the evaluation process (Darling-Hammond, Wise, & Pease, 1983; Stronge, 1997). To ensure the evaluation policy is clearly communicated, the available research suggests involving teachers in the design and implementation of the evaluation process (Kyriakides, Demetriou, & Charalambous, 2006).

EVALUATION INNOVATIONS

*By Angela Baber, Researcher for the Teacher Quality and Leadership Institute at the
Education Commission of the States*

Programs that evaluate teachers based on outcomes (such as teacher behavior in the classroom or student academic gains) rather than nonoutcome measures (such as certification and experience) are of increasing interest to policymakers and education leaders looking to tie teacher advancement to effectiveness. Two innovative systems for evaluating teachers are highlighted here. Minnesota Q Comp is a state-level, performance-pay program that includes an evaluation system and allows for district-level flexibility. Cincinnati Public Schools has established a comprehensive evaluation system used for teacher career advancement, but evaluation outcomes are not tied to performance pay.

Minnesota's Quality Compensation (Q Comp)

Quality Compensation, or Q Comp, is a performance-pay program adopted by the state of Minnesota. Participation in this program is not a state requirement; rather, districts apply to participate. Although the Minnesota Department of Education has established basic requirements that each district must address to be approved for funding, the program allows districts to establish their own evaluation standards.

Under Q Comp, every teacher must be evaluated multiple times each year using a comprehensive standards-based professional review system that utilizes input from a variety of sources, including instructional observations and standards-based assessments to determine student growth. The review system must be informed by scientifically based education research. Principals and peer reviewers such as master and mentor teachers conduct the teacher evaluations, and the evaluations must be one consideration for teacher bonuses (Minnesota Department of Education, 2007). In order to ensure fairness, all evaluators are required to use the same evaluation criteria.

Cincinnati Public Schools Teacher Evaluation System (TES)

Cincinnati Public Schools has implemented a comprehensive system called the Teacher Evaluation System (TES). The original plan was to have two phases of implementation; the second phase was intended to tie compensation to a teacher's TES ranking. However, this phase was voted down (Cincinnati Public Schools, n.d.). The current evaluation system uses annual evaluations to determine teacher movement on a traditional salary schedule and is based on 16 standards divided into four domains: planning and preparing for student learning, creating an environment for learning, teaching for learning, and professionalism.

Using a set of rubrics, administrators measure a teacher's performance against each of these standards. The results "place" teachers on one of five levels, and each increase in level is associated with a salary increase. If a teacher receives an evaluation that places him or her in a lower category, the teacher's salary increase is withheld and he or she must undergo a second comprehensive evaluation the following year.

For two of the TES domains—creating an environment for learning, and teaching for learning— evaluations are performed six times a year. Four of these evaluations are performed by a teacher from another school with subject-matter and grade-level expertise equivalent to the teacher being evaluated, and two are performed by school administrators. For the remaining two domains—planning and preparing for student learning, and professionalism—administrators evaluate teachers based on their portfolios including units and lesson plans, attendance records, student work, family contact logs, and documentation of professional development activities.

APPLICATION OF EVALUATION RESULTS

As previously indicated, formative evaluation results can be used to guide professional development plans and improve teacher practice. Using evaluation results to inform professional development empowers teachers to self-direct their growth (Nolan & Hoover, 2005) and encourages learning embedded in daily classroom practice.

Reality: Despite research and expert opinion on the value of aligning evaluations to professional development plans and the few examples in the field (see pages 14–15 of the Policy Options section for Iowa and Tennessee programs), the reality is that most teacher evaluations are summative in nature; therefore, most are commonly used to determine teacher employment status and personnel decisions, especially for nontenured teachers (e.g., Brandt et al., 2007). While summative evaluations are necessary, without formative feedback, teachers have little formal guidance to inform investments in professional development. The use of summative evaluation suggests that "school districts evaluate teachers simply because the law mandates evaluation, rather than as a way to guide staff development or improve instructional quality" (Zerger, 1988, p. 509). This compliance attitude toward teacher evaluation leads to inadequate allocation of the time and resources necessary to ensure effective evaluations (Zerger, 1988).

Furthermore, teacher evaluations that are a one-time snapshot of the teacher's practices make evaluators hesitant to be critical of the teacher. In unionized settings, the evaluators also may hesitate to act on evaluation results because of the legal cost associated with a potential grievance procedure (Bridges, 1992; Haefele, 1993; VanSciver, 1990). A grievance is most often filed when a teacher believes management is wrongfully seeking a dismissal or when a teacher is given a negative evaluation that he or she does not believe is warranted.

Recommended: Research convincingly demonstrates that when certain instructional strategies are implemented appropriately, they can increase student achievement (e.g., Marzano, Pickering, & Pollock, 2001). Teachers also have consistently reported the desire for feedback on how well or poorly they are implementing instructional strategies and delivering critical content. As such, it seems logical to assume that teacher evaluation results can provide teachers with the first step toward improving their instructional practices. Once communicated, the evaluation results should drive the individualized professional development opportunities that are made available to each teacher.

POLICY OPTIONS

The ways in which teacher performance and effectiveness are assessed have garnered increased interest and political visibility in recent years. A recent overview of several national, state, and local evaluation systems conducted by Education Sector (Toch & Rothman, 2008) eloquently highlights the urgent need for education policymakers to address the inadequate conduct of teacher evaluations and to emphasize their potential for teacher and school improvement. The approach of the National Comprehensive Center for Teacher Quality to this issue was to first deconstruct the many individual tools and instruments used as part of an evaluation process and to ground the following recommendations in available research, including the recent REL Midwest study. As a result, this Research and Policy Brief highlights the gap between current and recommended practices for evaluating teacher performance. Fortunately, however, there are many practical ways in which current teacher evaluation systems can be improved. Following are options for improving the use of teacher evaluation that states and districts might consider as levers in their ongoing efforts to improve student learning.

STATE POLICY OPTIONS

- **Create a statewide committee—composed of teachers, state and local teachers union or collective bargaining representatives, principals, and district administrators—to consider some of the challenges in designing and implementing an evaluation system and to recommend improvements in both evaluation implementation and the application of its results.** Influencing policymakers' attempts to improve teacher evaluation systems are state and district working relationships with teachers and teachers unions. By establishing a statewide committee with members representing all levels within the education system, it may be possible to start

a discussion about how best to measure teacher performance so all students may benefit from teachers' professional growth.

- **Develop a statewide bank of validated and reliable evaluation instruments, and advise districts to use multiple data sources.** States may want to consider creating a Web-based resource center that contains links to teacher evaluation instruments that have been shown to consistently work in local settings, including rubrics for analyzing classroom observations, scoring teacher portfolios, and reviewing student work and teacher contributions. States could partner with the Regional Educational Laboratory Network (http://ies.ed.gov/ncee/edlabs/) to identify valid and reliable instruments in their regions. Even a successfully validated evaluation instrument has limits to what it can measure. For that reason, it is advisable to consider the use of multiple measures of teacher performance.

- **Provide incentives and support for pilot programs in which groups of districts systematically test teacher evaluation measures.** Pilot programs create opportunities for states to identify which evaluation measures work best for informing and improving teacher practice, and in which contexts. Before a state launches an experimental program at the local level, however, it must ensure that the initiative is informed by a qualitative review of existing state evaluation policies and the extent to which they support or inhibit effective local teacher evaluation practices.

Field Example: During the 2001 Iowa legislative session, the Student Achievement and Teacher Quality (SATQ) program was established through Iowa Senate File 476 (2001). SATQ included the first part of a new four-level career ladder where advancement was determined on the basis

of teacher skills and knowledge—not experience and degrees. While the 2001 SATQ effort was a bold move to tie pay to performance, rewards were based on team-based effort rather than individual teacher performance. Teachers received monetary awards on top of their normal salaries. In 2001, the Iowa Legislature subsequently created the Teacher Pay-for-Performance (PFP) Commission to "design and implement a pay-for-performance program and provide a study relating to teacher and staff compensation structures containing pay-for-performance components" (Iowa Department of Management, 2007). The Teacher PFP Commission was formed to build on the 2001 legislation and examine options for tying individual teacher performance to teacher pay. It commenced its teacher pay-for-performance pilot program in July 2007, with up to 10 participating school districts (Iowa House File 2792, 2006).

• **Recommend that leadership preparation programs include knowledge of approaches to teacher evaluation as a core competency required for licensure.** One of the greatest challenges facing the consistent application of teacher evaluation practices is the paucity of trained and knowledgeable evaluators. Lack of training leads to the misuse of the evaluation instruments, the misinterpretation of results, and ultimately the lack of overall utility of the results for improving the performance of teachers. While the majority of evaluations are conducted by administrators, even in instances where a team-based approach to teacher evaluation is used, school leaders would benefit from a better understanding of how to conduct effective evaluations. Leadership preparation programs should encourage aspiring principals to learn how to record teaching practices, determine a teacher's impact on students, and use the results to align individual professional development opportunities for teachers with best practices (Goldrick, 2002).

Implementation Guideline: As part of the coursework that focuses on supervision and assessment, aspiring principals should be introduced to different evaluation measures and instruments, such as those presented in this brief, and learn to analyze and interpret student performance data in relation to teacher performance. During field experiences, principal candidates would observe the evaluation process and report on any improvements they would make to the instruments and processes. *Note:* Independent of who conducts the evaluation, all evaluators should be trained on the evaluation instruments and methods. Part of the training should focus on ensuring interrater reliability (i.e., all evaluators come to the same conclusion after using the same rubric for the same teacher).

• **Make teacher evaluation matter.** Even if an evaluation system is well designed (and perceived to be so by teachers), intrinsic motivation alone will not induce teachers, their peers, and their supervisors to take evaluation seriously. However, creating a sense of accountability relating to the results may make evaluation matter. A good starting place is to consistently connect evaluation results to investments in teacher professional development. Teachers may feel empowered and supported by the evaluation process if they see that it is designed to sustain their growth.

Field Example: Efforts to align evaluations with professional development are occurring in Tennessee. With assistance from administrators, teachers in Tennessee create professional development plans that focus on their individual growth in a specified performance standard. Teachers are then evaluated on a given set of goals addressing certain development needs around the performance standard (Tennessee Department of Education, 1998). In Iowa, the 2001 SATQ legislation required districts to develop an individual career development plan—in cooperation with

the teacher's supervisor—that is aligned with the Iowa Teaching Standards, the appropriate student achievement goals of the district, and the teacher's individual needs (Keystone Area Education Agency, 2003–04).

- **If using students' standardized test scores as part of the teacher evaluation process, provide technical support to districts.** Districts will grapple not only with the issue of data infrastructure but also with comprehending the technical and statistical procedures that build the indexes and other constructs that allow the comparisons.

- **Require state-funded district pilot programs to demonstrate how the application of evaluation results not only will be tied to monetary rewards but also will align with local and individual priorities for job-embedded professional development.** Strong evaluation systems indicate the type of professional development that would be most beneficial for teachers. Teacher survey data consistently report that teachers prefer opportunities to engage in high-quality professional learning over monetary incentives (Rochkind, Immerwahr, Ott, & Johnson, 2007). States may consider requiring districts to develop individualized professional development plans for all teachers based on individualized evaluation results to systematically improve students' learning opportunities. Linking professional development plans with practical, job-embedded opportunities will assist teachers in working toward their professional goals while working with their students (rather than hypothetically discussing strategies in a professional development workshop).

Local Policy Options

- **Enable experienced and exemplary teachers to serve as evaluators.** Across districts, the evaluator-to-teacher ratio may contribute to the brevity and infrequency with which evaluations are conducted. As previously stated, administrators most often are responsible for evaluating teachers; however, a common criticism of administrators as evaluators is that they are disconnected from the day-to-day intricacies of delivering and adjusting instruction to meet the needs of students in a particular classroom. The role of a peer evaluator could provide a leadership opportunity for exemplary teachers seeking to expand their career experiences beyond the classroom and reduce the burden placed on principals.

- **Increase the frequency of formative evaluations**. Painting a more accurate picture of teacher performance requires the frequent use of formative evaluation instruments. Through frequent evaluation activities, evaluators gain an understanding of the dynamics in a particular classroom and how certain instructional strategies may work better under certain conditions. Formative evaluations that occur periodically throughout the school year may provide ongoing and critical feedback to teachers about their practices and inform administrators about buildingwide issues. Also, formative assessments may capture teachers' improvements over time more accurately than infrequent summative evaluations.

- **Consider using more frequent evaluations to inform the professional growth of all teachers.** The frequency and intensity of evaluation activities often is determined by teacher tenure (i.e., those with more experience are evaluated less frequently). Tenured teachers often receive less feedback about their teaching practices, which could hamper their professional

growth. Like other professionals, all teachers should receive feedback more than once a year; however, the exact frequency with which one is evaluated may be determined by individual needs. The need for more frequent or extensive evaluations can be determined on the grounds of instructional needs and strengths rather than tenure.

- **Use evaluation results to inform the professional development opportunities that districts and schools make available to teachers.** Teacher evaluation systems, if constructed and used appropriately, could reveal teachers' instructional strengths and areas in need of growth over time. Armed with this information, teachers could set their individualized professional goals based on evaluation feedback. Similarly, a collective picture of the staff's professional needs could guide the decisions that districts and schools make regarding investments in professional development.

  Field Example: Vaughn Elementary School in the Los Angeles Unified School District uses evaluation results to suggest professional development opportunities to teachers in the areas identified by their evaluations as needing more attention (Gallagher, 2004). Most Vaughn teachers reported that their evaluation system focuses on improving instruction, increasing student achievement, and helping the teaching staff develop additional skills (Kellor, 2005).

- **Develop a review process and communication plan to gauge teacher and administrator perceptions and concerns of the evaluation system and revise the system as necessary.** To ensure that the evaluation system is responsive to teachers' needs and that it is producing the expected experiences and outcomes for evaluators, teachers, and schools, ongoing feedback about the evaluation process,

procedures, and measures should be collected in a systemic manner. In addition to collecting data, a format for dialogue with teachers and evaluators about their concerns and suggestions for improvements should be in place. These steps ensure that the evaluation remains a dynamic system that continues to be valued over time.

Implementation Guideline: Part of the district's data management system could include a database of teachers' and administrators' perceptions of the evaluation system. Questions about one's perceptions of the evaluation system could be embedded in an existing annual or biannual survey that seeks to understand educators' views on various district initiatives and processes. A committee of researchers, administrators, and teachers could be formed to review the survey results and determine what revisions could be made to respond to perception data.

CONCLUSION

Transforming teacher evaluation systems into mechanisms for improving student learning is a challenge with deep roots in the national debate about teacher quality and how to measure and reward teacher excellence. To inform evaluation practices, future research should explore (1) the role of union contracts in teacher evaluations; (2) the role of state policy in directing teacher evaluation at the district level; (3) how state education departments support the teacher evaluation process; (4) variations in state language and policy specificity and how these issues impact teacher evaluation at the local level; (5) the influence of district policy on the evaluation of beginning (nontenured), experienced (tenured), and unsatisfactory teachers; (6) the impact of evaluation models and practice on teacher effectiveness; and (7) the relationship between the number of teachers assigned to an evaluator and the impact of that number on the reliability and validity of their evaluations. (See the sidebar below concerning future research.)

However, without a careful review and inclusive dialogue at the state and local levels about how to improve approaches to teacher evaluation, opportunities to truly influence changes in teacher quality are mostly empty promises. If the education system is unable to provide formative and summative feedback to its teachers, not only does it fail teachers; in the end, it also fails children. Given the overwhelming evidence that good teachers have the greatest impact on positive student outcomes, supporting their ongoing growth and development ought to be a priority in education. Without the appropriate assessments to identify problems and recognize excellence, investments in teacher development are disconnected from school and district goals for improvement. This Research and Policy Brief provides information to encourage states and districts to assess the appropriateness and effectiveness of their teacher evaluation systems.

RESEARCH FINDINGS ON THE HORIZON

As a follow-up to its descriptive study of districts' teacher evaluation policies, REL Midwest is currently examining the alignment between district and state policy concerning teacher evaluation. In the current policy environment, the state-level priority to improve teacher quality—combined with pressure to improve achievement for all students—places the issue of teacher evaluation at the center (Goldrick, 2002; Gordon et al., 2006). Clearly, states will face major challenges in enacting policies, codes, rules, and regulations to guide the creation of local teacher evaluation systems that are aligned to clear teacher performance standards, incorporate multiple data sources to inform comprehensive teacher evaluations, include measures for accurately factoring student achievement growth into evaluations, and emphasize the use of results to inform individualized teacher professional development plans. Although Midwestern states have begun to address these challenges, it is still unclear how teacher evaluation policy and practices vary at the state level; also unclear is the extent to which state policies align and support district practice. The REL Midwest study currently under way will describe how the Midwestern states are dealing with these issues and perhaps serve as a guide for supporting more effective state and district teacher evaluation policies.

REFERENCES

Attinello, J. R., Lare, D. W., & Source, F. (2006). The value of teacher portfolios for evaluation and growth. *NASSP Bulletin, 90*(2), 132–152.

Andrejko, L. (1998). The case for the teacher portfolio: Evaluation tool carries a wealth of professional information. *Journal of Staff Development, 19*(4). Retrieved February 25, 2008, from http://www.nsdc.org/library/publications/jsd/andrejko194.cfm

Barrett, J. (1986). The evaluation of teachers. *ERIC Digest.* Washington, DC: ERIC Clearinghouse on Teacher Education (ERIC Identifier: ED278657). Retrieved February 25, 2008, from http://www.ericdigests.org/pre-925/evaluation.htm

Berry, B. (2007, March). *Connecting teacher and student data: Benefits, challenges, and lessons learned.* Presentation for the Data Quality Issue meeting, Washington, DC. Retrieved February 25, 2008, from http://www.dataqualitycampaign.org/files/Presentations-DQC_Quaterly_Meeting_Connecting_Teacher_Student_Data_031207.pdf

Blunk, M. (2007, April). *The QMI: Results from validation and scale-building.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Brandt, C., Mathers, C., Oliva, M., Brown-Sims, M., & Hess, J. (2007). *Examining district guidance to schools on teacher evaluation policies in the Midwest Region* (Issues & Answers Report, REL 2007–No. 030). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved February 25, 2008, from http://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_2007030.pdf

Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models.* Princeton, NJ: Educational Testing Service. Retrieved February 27, 2008, from http://www.ets.org/Media/Research/pdf/PICVAM.pdf.

Bridges, E. (1992). *The incompetent teacher: Managerial responses.* Philadelphia: Falmer Press.

Cincinnati Public Schools. (n.d.). *Teacher evaluation.* Retrieved February 25, 2008, from http://www.cps-k12.org/employment/tchreval/tchreval.htm

Cronin, L., & Capie, W. (1986, April). *The influence of daily variation in teacher performance on the reliability and validity of assessment data.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching.* Alexandria, VA: Association for Supervision and Curriculum Development & Educational Testing Service.

Danielson, C., & McGreal, T. L. (2000). *Teacher evaluation to enhance professional practice.* Alexandria, VA: Association for Supervision and Curriculum Development & Educational Testing Service.

Darling-Hammond, L., Wise, A. E., & Pease, S. R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research, 53*(3), 285–328.

Denner, P. R., Miller, T. L., Newsome, J. D., & Birdsong, J. R. (2002). Generalizability and validity of the use of a case analysis assessment to make visible the quality of teacher candidates. *Journal of Personnel Evaluation in Education, 16*(3), 153–174.

Denner, P. R., Salzman, S. A., & Bangert, A. W. (2001). Linking teacher assessment to student performance: A benchmarking, generalizability, and validity study of the use of teacher work samples. *Journal of Personnel Evaluation in Education, 15*(4), 287–307.

Ellett, C. D., & Garland, J. (1987). Teacher evaluation practices in our largest school districts: Are they measuring up to "state-of-the-art" systems? *Journal of Personnel Evaluation in Education, 1*(1), 69–92.

Gallagher, H. A. (2004). Vaughn Elementary's innovative teacher evaluation system: Are teacher evaluation scores related to growth in student achievement? *Peabody Journal of Education, 79*(4): 79–107.

Goldstein, J., & Noguera, P. A. (2006, March). A thoughtful approach to teacher evaluation. *Educational Leadership, 63*(6), 31–37.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved February 25, 2008, from http://www.ncctq.org/publications/LinkBetweenTQandStudentOutcomes.pdf

Goldrick, L. (2002). *Improving teacher evaluation to improve teaching quality* (Issue Brief). Washington, DC: National Governors Association, Center for Best Practices. Retrieved February 25, 2008, from http://www.nga.org/Files/pdf/1202improvingteacheval.pdf

Gordon, R., Kane, T., & Staiger, D. O. (2006). *Identifying effective teachers using performance on the job* (Policy Brief No. 2006-01). Washington, DC: The Brookings Institution.

Haefele, D. L. (1993). Evaluating teachers: a call for change. *Journal of Personnel Evaluation in Education, 7*(1), 21–31.

Iowa Department of Management. (2007). *Pay for Performance Commission.* Retrieved February 25, 2008, from http://www.dom.state.ia.us/pfp_commission/index.html

Iowa House File 2792. (2006). Retrieved February 25, 2008, from http://coolice.legis.state.ia.us/Cool-ICE/default.asp?Category=billinfo&Service=Billbook&frame=1&GA=81&hbill=HF2792

Iowa Senate File 476. (2001). Retrieved February 25, 2008, from http://www.legis.state.ia.us/GA/79GA/Legislation/SF/00400/SF00476/Current.html

Kellor, E. M. (2005). Catching up with the Vaughn express: Six years of standards-based teacher evaluation and performance pay. *Education Policy Analysis Archives, 13*(7). Retrieved February 25, 2008, from http://epaa.asu.edu/epaa/v13n7

Keystone Area Education Agency. (2003–04). *Summary of the Iowa professional development model.* Retrieved February 25, 2008, from http://www.aea1.k12.ia.us/profdevelopment/pdsummary.html

Kyriakides, L., Demetriou, D., & Charalambous, C. (2006). Generating criteria for evaluating teachers through teacher effectiveness research. *Educational Research, 48*(1), 1–20.

Loup, K. S., Garland, J. S., Ellett, C. D., & Rugutt, J. K. (1996). Ten years later: Findings from a replication of a study of teacher evaluation practices in our 100 largest school districts. *Journal of Personnel Evaluation in Education, 10*(3), 203–226.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, 79*(4), 33–53.

Minnesota Department of Education. (2007). *Quality compensation for teachers (Q Comp).* Retrieved February 25, 2008, from http://education.state.mn.us/MDE/Teacher_Support/QComp/index.html

Mujis, D. (2006). Measuring teacher effectiveness: Some methodological reflections. *Educational Research and Evaluation, 12*(1), 53–74.

National Council on Teacher Quality. (2006). *Teacher rules, roles and rights.* Retrieved February 25, 2008, from http://www.nctq.org/tr3/

Nolan, J., Jr., & Hoover, L. A. (2005). *Teacher supervision and evaluation: Theory into practice.* New York: Wiley.

Parkes J., & Stevens, J. (2000). *How professional development, teacher satisfaction, and plans to remain in teaching are related: Some policy implications* (Research Report 2000-1). Albuquerque, NM: Albuquerque Public Schools, University of New Mexico, & Albuquerque Teacher Federation Partnership. Retrieved February 25, 2008, from http://citeseer.ist.psu.edu/cache/papers/cs/23189/http:zSzzSzwww.unm.eduzSz~parkeszSzretention.pdf/how-professional-development-teacher.pdf

Peterson, P. L., & Comeaux, M. A. (1990). Evaluating the systems: Teachers' perspectives on teacher evaluation. *Educational Evaluation and Policy Analysis, 12*(1), 3–24.

Price, J. and Schwabacher, S. (1993). *The multiple forms of evidence study: Assessing reading through student work samples, teacher observations, and tests.* New York: National Center for Restructuring Education, Schools, and Teaching. Retrieved February 25, 2008, from http://www.eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED372366&ERICExtSearch_SearchType_0=no&accno=ED372366

Rochkind, J., Immerwahr, J., Ott, A., & Johnson, J. (2007). Getting started: A survey of new public school teachers on their training and first months on the job. In C. A. Dwyer (Ed.), *America's challenge: Effective teachers for at-risk schools and students* (pp. 89–104). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved February 25, 2008, from http://www.ncctq.org/publications/NCCTQBiennialChapters/NCCTQ%20Biennial%20Report_Ch6.pdf

Schon, D. A. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic Books.

Shannon, D. M. (1991, February). *Teacher evaluation: A functional approach.* Paper presented at the annual meeting of the Eastern Educational Research Association, Boston.

Shavelson, R. J., Webb, N. M., & Burstein, L. (1986). The measurement of teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 55–91). New York: Simon and Schuster MacMillan.

Shinkfield, A. J., & Stufflebeam, D. L. (1995). *Teacher evaluation: Guide to effective practice.* Boston: Springer.

Stiggans, R. J., & Duke, D. L. (1988). *The case for commitment to teacher growth: Research on teacher evaluation.* Albany, NY: State University of New York Press.

Stronge, J. H. (1997). Improving schools through teacher evaluation. In J. H. Stronge (Ed.), *Evaluating teaching: A guide to current thinking and best practice* (pp. 1–23). Thousand Oaks, CA: Corwin Press.

Stronge, J. H. (2007). Planning and organizing for instruction. In J. H. Stronge, *Qualities of effective teachers* (2nd ed., Chapter 4). Alexandria, VA: Association for Supervision and Curriculum Development. Retrieved February 25, 2008, from http://www.ascd.org/portal/site/ascd/template.chapter/menuitem.b71d101a2f7c208cdeb3ffdb62108a0c/?chapterMgmtId=fa8db5f04e130110VgnVCM1000003d01a8c0RCRD

Sweeney, J., & Manatt, R. (1986). Teacher evaluation. In R. Berk (Ed.), *Performance assessment: Methods and applications* (pp. 446–468). Baltimore: John Hopkins University Press.

Tennessee Department of Education. (1998). *Professional growth plan.* Retrieved February 25, 2008, from http://www.state.tn.us/education/frameval/pgp.pdf

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education.* Washington, DC: Education Sector. Retrieved February 25, 2008, from http://www.educationsector.org/usr_doc/RushToJudgment_ES_Jan08.pdf

Tucker, P. D., Stronge, J. H., & Gareis, C. R. (2002). *Handbook on teacher portfolios for evaluation and professional development.* Larchmont, NY: Eye on Education.

Uhlenbeck, A. M., Verloop, N., & Beijaard, D. (2002). Requirements for an assessment procedure for beginning teachers: Implications from recent theories on teaching and assessment. *Teachers College Record, 104*(2), 242–272.

VanSciver, J. H. (1990). Teacher dismissals. *Phi Delta Kappan, 72*(4), 318–319.

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Case studies for teacher evaluation: A study of effective practices.* Santa Monica, CA: RAND Corporation. Retrieved February 25, 2008, from http://www.rand.org/pubs/notes/2007/N2133.pdf

Zerger, K. L. (1988). Teacher evaluation and collective bargaining: A union perspective. *Journal of Law and Education, 17*(3), 507–525.

About the National Comprehensive Center for Teacher Quality

The National Comprehensive Center for Teacher Quality was launched on October 2, 2005, after Learning Point Associates and its partners—Education Commission of the States, ETS, and Vanderbilt University—entered into a five-year cooperative agreement with the U.S. Department of Education to operate the teacher quality content center.

It is a part of the U.S. Department of Education's Comprehensive Centers program, which includes 16 regional comprehensive centers that provide technical assistance to states within a specified boundary and five national content centers that provide expert assistance to benefit states and districts nationwide on key issues related to the goals of the No Child Left Behind Act.

NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

1100 17th Street NW, Suite 500
Washington, DC 20036-4632
877-322-8700 • 202-223-6690
www.ncctq.org

2131_02/08

LEARNING POINT Associates

Education Commission of the States

ETS

VANDERBILT